

10/506651
DT15 Rec'd PCT/PTO 02 SEP 2004

SEQUENCE LISTING

SEQ ID NO:1 full length nucleic acid 89.6 gp160 Env sequence

```

atgagagt gaaggagatc aggaagaatt ggcagcactt gagagggggc
atcttgctcc ttgggatggt gatgatctgt agtgctgcaa aagaaaaaac
gtgggtcaca atctattatg ggggtacctgt gtggagagaa gcaaccacca
ctctattttg tgcatacagat gctaaagcct atgatacaga ggtacataat
gtttgggcca cacatgcctg tgtacccaca gaccccaacc cacaagaagt
agtattggga aatgtgacag aaaattttta catgtggaaa aataacatgg
tagatcagat gcatgaggat ataatacagt tatgggatga aagcctaaag
ccatgtgtaa aattaacccc actctgtgtt actttaaatt gcactaattt
gaatatcact aagaatacta ctaatccac tagtagcagc tggggaatga
tggagaaagg agaaataaaa aattgctctt tctatatcac cacaagcata
agaaataagg taaagaaaga atatgcactt tttaatagac ttgatgtagt
accaatagaa aatactaata atactaagta taggttaata agttgtaaca
cctcagtcac tacacaggcc tgtccaaagg taccctttca gccaatcccc
atacattatt gtgtcccggc tgggtttgcg atgctaaagt gtaacaataa
gacattcaat ggatcaggac catgcacaaa tgtcagcaca gtacaatgta
cacatggaat taggccagt gtgtcaactc aactgctgtt aaatggcagt
ctagcagaag aagacatagt aattagatct gaaaatttca cagacaatgc
taaaaccata atagtacagc taaatgaatc tgtagtaatt aattgtacaa
gacccaacaa caatacaaga agaaggttat ctataggacc agggagagca
ttttatgcaa gaagaaacat aataggagat ataagacaag cacattgtaa
cattagtaga gcaaaatgga ataacacttt acaacagata gttataaaat
taagagaaaa attttaggaat aaaacaatag cctttaatca atcctcagga
ggggaccag aaattgtaat gcacagtttt aattgtggag ggggaattttt
ctactgtaat acagcacaaac tgtttaatag tacttggaat gttactggag
ggacaaatgg cactgaagga aatgacataa tcacactcca atgcagaata
aaacaaatta taaatatgtg gcagaaagta ggaaaagcaa tgtatgcccc
tcccatcaca ggacaaatta gatgttcac aaatattaca gggctgctac
taacaagaga tggaggtaat agtactgaga ctgagactga gatcttcaga
cctggaggag gagatatgag ggacaatttg agaagtgaat tatataaata
taaagtagta agaattgaac caataggagt agcaccacc agggcaaaga
gaagaacagt gcaaagagaa aaaagagcag tgggaatagg agctgtgttc
cttgggttct tgggagcagc aggaagcact atgggcgcag cgtcagtgac
gctgacggta caggccaggc tattattgtc tggatatagt cagcagcaga
acaatctgct gagggctatt gaggcgcaac agcatatgtt gcaactcaca
gtctggggca tcaagcagct ccaggcaaga gtcctggctc tggaaagata
cctaagggat caacagctca tgggaatttg gggttgctct ggaaaactca
tttgcaccac ttctgtgcct tggaatgtta gttggagtaa taaatctgtg
gatgatattt ggaataacat gacctggatg gagtgggaaa gagaaattga
caattacaca gactatatat atgacttact tgaaaaatcg caaacccaac
aagaaaagaa tgaaaaagaa ttattggaat tggataaatg ggcaagtttg
tggaattggt ttgacataac aaactggctg tggatatata gattattcat
aatgatagta ggaggcttga taggtttaag aatagttttt gctgtacttt
ctatagtaaa tagagttagg cagggatatt caccattatc gtttcagacc

```

ctcctcccag cctcgagggg acccgacagg cccgaaggaa cagaagaaga
 aggtggagag agagacagag acagatccgg tccattagtg aacggattct
 tggcactttt ctgggtcgat ttgaggaacc tgtgcctctt cctctaccac
 ctcttgagaa acttactctt gattgtaacg aggattgtgg aacttctggg
 acgcaggggg tgggaagccc tcaaattatt gtggaatctc ctgcaatatt
 ggagccagga actaaagaat agtgctgtta gcttgctcaa cgccacagcc
 atagcagtag ctgagggggac agatagggtt ataaaaatag tacaagagc
 ttgtagagct attcgcaaca tacctacaag aatcagacag ggcttggaaa
 gggctttgct ataa

SEQ ID NO:2 Amino acid sequence of 89.6 gp160

1 MRVKEIRKNW QHLRGGILL GMLMICSAAK EKTWVTIYYG VPVWREATTT
 51 LFCASDAKAY DTEVHNWAT HACVPTDPNP QEVVLGNVTE NFNMWKNNMV
 101 DQMHEIISL WDESLKPCVK LTPLCVTLNC TNLNITKNTT NPTSSSWGMM
 151 EKGEIKNCSF YITTSIRNKV KKEYALFNRL DVVPIENTNN TKYRLISCNT
 201 SVITQACPKV SFQPIPIHYC VPAGFAMLKC NNKTFNGSGP CTNVSTVQCT
 251 HGIRPVVSTQ LLLNGSLAEE DIVIRSENF DNAKTIIVQL NESVVINCTR
 301 PNNNTRRRLS IGPGRFYAR RNIIGDIRQA HCNISRAKWN NTLQQIVIKL
 351 REKFRNKTI FNOSSGGDPE IVMHSEFNCGG EFFYCNTAQL FNSTWNVTTG
 401 TNGTEGNDII TLQCRKQII NMWQKVGKAM YAPPITGQIR CSSNITGLLL
 451 TRDGGNSTET ETEIFRPGGG DMRDNWRSEL YKYKVVRIEP IGVAPTRAKR
 501 RTVQREKRAV GIGAVFLGFL GAAGSTMGA SVTLTVQARL LLSGIVQQQN
 551 NLLRAIEAQQ HMLQLTVWGI KQLQARVLAL ERYLRDQQLM GIWGC SGKLI
 601 CTTSPVWNVS WSNKSVDIWN NMMTWMEWER EIDNYTDYIY DLLEKSQTQQ
 651 EKNEKELLEL DKWASLWNWF DITNWLWYIR LFIMIVGGLI GLRIVFAVLS
 701 IVNRVRQGYS PLSFQTLPA SRGPDRPEGT EEEGGERDRD RSGPLVNGFL
 751 ALFWVDLRNL CLFLYHLLRN LLLIVTRIVE LLGRRGWEAL KYWWNLLQYW
 801 SQELKNSAVS LLNATAIAVA EGTDRVIKIV QRACRAIRNI PTRIRQGLER
 851 ALL*

SEQ ID NO:3 89.6 gp140 Nucleic Acid sequence

atgagagt gaaggagatc aggaagaatt ggcagcactt gagagggggg
 atcttgctcc ttgggatgtt gatgatctgt agtgctgcaa aagaaaaaac
 gtgggtcaca atctattatg gggtagctgt gtggagagaa gcaaccacca
 ctctattttt tgcatacagat gctaaagcct atgatacaga ggtacataat
 gtttgggcca cacatgcctg tgtaccacca gacccaacc cacaagaagt
 agtattggga aatgtgacag aaaattttta catgtggaaa aataacatgg
 tagatcagat gcatgaggat ataatacgtt tatgggatga aagcctaaag
 ccatgtgtaa aattaacccc actctgtgtt actttaaatt gcactaat
 gaatatcact agaataacta ctaatccac tagtagcagc tggggaatga
 tggagaaagg agaaataaaa aattgctctt tctatatcac cacaagcata
 agaaataagg taaagaaaga atatgcactt ttaataagac ttgatgtagt
 accaatagaa aataactaata atactaagta taggttaata agttgtaaca
 cctcagtcac tacacaggcc tgtccaaagg tatcctttca gccaatccc
 atacattatt gtgtcccggc tgggtttgag atgctaaagt gtaacaataa
 gacattcaat ggatcaggac catgcacaaa tgtcagcaca gtacaatgta
 cacatggaat taggccagt gtgtcaactc aactgctgtt aatggcagt
 ctagcagaag aagacatagt aattagatct gaaaatttca cagacaatgc
 taaaaccata atagtagcag taaatgaatc ttagtagtaatt aattgtacaa

```

gacccaacaa caatacaaga agaagggttat ctataggacc agggagagca
ttttatgcaa gaagaaacat aataggagat ataagacaag cacattgtaa
cattagtaga gcaaaatgga ataacacttt acaacagata gttataaaat
taagagaaaa atttaggaat aaaacaatag cctttaatca atcctcagga
ggggacccag aaattgtaat gcacagtttt aattgtggag gggaattttt
ctactgtaat acagcacaac tgtttaatag tacttggaat gttactggag
ggacaaatgg cactgaagga aatgacataa tcacactcca atgcagaata
aaacaaatta taaatatgtg gcagaaagta ggaaaagcaa tgtatgcccc
tcccatcaca ggacaaatta gatgttcac aaatattaca gggctgctac
taacaagaga tggaggtaat agtactgaga ctgagactga gatcttcaga
cctggaggag gagatatgag ggacaattgg agaagtgaat tatataaata
taaagtagta agaattgaac caataggagt agcaccacc agggcaaaga
gaagaacagt gcaaagagaa aaaagagcag tgggaatagg agctgtgttc
cttgggttct tgggagcagc aggaagcact atgggagcag cgtcagtgc
gctgacggta caggccaggc tattattgtc tggatatagt cagcagcaga
acaatctgct gagggctatt gaggcgcaac agcatatgtt gcaactcaca
gtctggggca tcaagcagct ccaggcaaga gtcctggctc tggaaagata
cctaagggat caacagctca tgggaatttg gggttgctct ggaaaactca
tttgcaccac ttctgtgcct tggaatgtta gttggagtaa taaatctgtg
gatgatattt ggaataacat gacctggatg gagtgggaaa gagaaattga
caattacaca gactatatat atgacttact tgaaaaatcg caaacccaac
aagaaaagaa tgaaaaagaa ttattggaat tggataaatg ggcaagtttg
tgggaattgg ttgacataac aaactggctg tggatatataa gattattcat
aat

```

SEQ ID NO:4 Amino acid sequence of 89.6 gp140

```

1  MRVKEIRKNW QHLRGGILLL GMLMCSAAK EKTWVTIYYG VPVWREATTT
51  LFCASDAKAY DTEVHNWVAT HACVPTDPNP QEVVLGNVTE NFNMWKNNMV
101 DQMHEDIISL WDESLKPCVK LTPLCVTLNC TNLNITKN TT NPTSSSWGMM
151 EKGEIKNCSF YITTSIRNKV KKEYALFNRL DVVPIENTNN TKYRLISCNT
201 SVITQACPKV SFQPIPIHYC VPAGFAMLC NNNKTFNGSGP CTNVSTVQCT
251 HGIRPVVSTQ LLLNGSLAEE DIVIRSENFT DNAKTIIVQL NESVVINCTR
301 PNNNTRRRLS IGPGRIFYAR RNIIGDIRQA HCNISRAKWN NTLQQIVIKL
351 REKFRNKTIA FNQSSGGDPE IVMHSFNCGG EFFYCNTAQL FNSTWNV TGG
401 TNGTEGNDII TLQCRIKQII NMWQKVGKAM YAPPITGQIR CSSNITGLLL
451 TRDGGNSTET ETEIFRPGGG DMRDNWRSEL YKYKVVRIEP IGVAPTRAKR
501 RTVQREKRAV GIGAVFLGFL GAAGSTMGA SVTLTVQARL LLSGIVQQQN
551 NLLRAIEAQQ HMLQLTVWGI KQLQARVLAL ERYLRDQQLM GIWGC SGKLI
601 CTTSPVWNVS WSNKSVDDIW NNMTWMEWER EIDNYTDYIY DLLEKSQTQQ
651 EKNEKELLEL DKWASLW*

```

SEQ ID NO:5: Nucleic acid sequence of env 89.6 indicating cleavage site and start of the gp41 region.

7747

CAA AGA GAA AAA AGA GCA GTG GGA ATA

SEQ ID NO:6 Amino acid sequence of env 89.6 indicating cleavage site and start of the gp41 region.

| | | | | | | | |
|-----|-----|-----|-----|-----|---|---|---|
| R | E | K | R | A | V | G | I |
| 505 | 506 | 507 | 508 | 509 | | | |

SEQ ID NO:7 Amino acid sequence of gp140-26 (soluble, secreted protein)

```

1  KEKTWVTIYY GVPVWREATT TLFCASDAKA YDTEVHNGWA THACVATDPN
51  PQEVVLGNVT ENFNMWKNNM VDOMHEDIIS LLDESLKPCV KLTPLCVTLN
101 CTNLNITKNT TNPTSSSLGM MEKGEIKNCS FYITTSIRNK VKKEYALFNR
151 LDVVPIENTN NTKYRLISCN TSVITQACPK VFFQPIAIHY CVPAGFAMLK
201 CNNKTFNGSG PCTNVSTVPC THGIRPVVST QLLLNGLAE EDIVIRSGNF
251 TDNAKTIIVQ LNESVVINCT RPNNNTRRRL SIGPGRAFYA RRNIIGDIRQ
301 AHCNISRAKL NNTLQQIVIK LREKFRNKTI AFNQSSGGDP EIVMHSFNCG
351 GEFFYCNTAQ LFNSTLNV TG GTNGTEENDI ITLQCRIKQI INMWQKVGKA
401 MYAPPITGQI ICSSNITGLL LTRDGGNSTE TETEIFRPGG GDMRDNWRSE
451 LYKYKVVRIE PIGVAPTRAK RRTCQGGIDG ILOISGSGSG GSGOGSSSGG
501 AGGKGAVGIG AVFLGFLGAA GSTMGARSVT LTVQARLLLS GIVQQQNLL
551 RAIEAQQHML QLTWVGKQL QARVLALERY LRDQQLMGIW GCSGKLICTT
601 SVPWNVSWSN KSVDDIWNM TWELEREID NYTDYIYDLL EKSQTQQEK
651 EKELLELDKW ASLWKLV

```

SEQ ID NO:8 Amino acid sequence of gp140-15 (soluble, secreted protein)

```

1  KEKTWVTIYY GVPVWREATT TLFCASDAKA YDTEVHNGWA THACVATDPN
51  PQEVVLGNVT ENFNMWKNNM VDOMHEDIIS LLDESLKPCV KLTPLCVTLN
101 CTNLNITKNT TNPTSSSLGM MEKGEIKNCS FYITTSIRNK VKKEYALFNR
151 LDVVPIENTN NTKYRLISCN TSVITQACPK VFFQPIAIHY CVPAGFAMLK
201 CNNKTFNGSG PCTNVSTVPC THGIRPVVST QLLLNGLAE EDIVIRSGNF
251 TDNAKTIIVQ LNESVVINCT RPNNNTRRRL SIGPGRAFYA RRNIIGDIRQ
301 AHCNISRAKL NNTLQQIVIK LREKFRNKTI AFNQSSGGDP EIVMHSFNCG
351 GEFFYCNTAQ LFNSTLNV TG GTNGTEENDI ITLQCRIKQI INMWQKVGKA
401 MYAPPITGQI ICSSNITGLL LTRDGGNSTE TETEIFRPGG GDMRDNWRSE
451 LYKYKVVRIE PIGVAPTRAK RRTCQGGIDG ILOISSSGGA GCKGAVGIGA
501 VFLGFLGAAG STMGARSVT LTVQARLLLSG IVQQQNLLR AIEAQQHMLQ
551 LTVWGIKQL ARVLALERYL RDQQLMGIWG CSGKLICTTS VPWNVSWSNK
601 SVDDIWNMT WMELEREIDN YTDYIYDLE KSQTQQEKNE KELLELDKWA
651 SLWKLV

```

SEQ ID NO :9 Four Amino acid peptide linker sequence GILI

SEQ ID NO :10 Fifteen Amino acid peptide linker sequence GGIDGILQISSSGGA

SEQ ID NO :11 Twenty-six Amino acid peptide linker sequence

GGIDGILQISGSGSGGSGQGSSSGGA

**SEQ ID NO :12 Fifteen Amino acid peptide linker sequence
(GGSGG)₃**

**SEQ ID NO :13 Twenty Amino acid peptide linker sequence
(GGSGG)₄**

**SEQ ID NO :14 Twenty five Amino acid peptide linker sequence
(GGSGG)₅**

PA 3205456 v1